Mol Gen Genomics (2004) 272: 181–193
DOI 10.1007/s00438-004-1035-3

ORIGINAL PAPER

**L. D. Cooper · L. Marquez-Cedillo · J. Singh
A. K. Sturbaum · S. Zhang · V. Edwards · K. Johnson
A. Kleinhofs · S. Rangel · V. Carollo
P. Bregitzer · P. G. Lemaux · P. M. Hayes**

Mapping *Ds* insertions in barley using a sequence-based approach

Received: 19 December 2003 / Accepted: 11 June 2004 / Published online: 30 July 2004

Abstract A transposon tagging system, based upon maize *Ac*/*Ds* elements, was developed in barley (*Hordeum vulgare* subsp. *vulgare*). The long-term objective of this project is to identify a set of lines with *Ds* insertions dispersed throughout the genome as a comprehensive tool for gene discovery and reverse genetics. *Ac*TPase and *Ds-bar* elements were introduced into immature embryos of Golden Promise by biolistic transformation. Subsequent transposition and segregation of *Ds* away from *Ac*TPase and the original site of integration resulted in new lines, each containing a stabilized *Ds* element in a new location. The sequence of the genomic DNA flanking the *Ds* elements was obtained by inverse PCR and TAIL-PCR. Using a sequence-based mapping strategy, we determined the genome locations of the *Ds* insertions in 19 independent lines using primarily restriction digest-based assays of PCR-amplified single nucleotide polymorphisms and PCR-based assays of insertions or deletions.The proncipal strategy was to identify and map sequence polymorphisms in the regions corresponding to the flanking DNA using the Oregon Wolfe Barley mapping population. The mapping results obtained by the sequence-based approach were confirmed by RFLP analyses in four of the lines. In addition, cloned DNA sequences corresponding to the flanking DNA were used to assign map locations to Morex-derived genomic BAC library inserts, thus integrating genetic and physical maps of barley. BLAST search results indicate that the majority of the transposed *Ds* elements are found within predicted or known coding sequences. Transposon tagging in barley using *Ac*/*Ds* thus promises to provide a useful tool for studies on the functional genomics of the Triticeae.

Keywords Transposon tagging · Barley · Reverse genetics · *Ac*/*Ds* · Mapping

Communicated by M.-A. Grandbastien
The first three authors contributed equally to this work

Electronic Supplementary Material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s00438-004-1035-3

L. D. Cooper · L. Marquez-Cedillo · P. M. Hayes (✉)
Department of Crop and Soil Science,
Oregon State University, Corvallis,
OR 97331, USA
E-mail: Patrick.M.Hayes@oregonstate.edu
Tel.: +1-541-7375878
Fax: +1-541-7370909

J. Singh · S. Zhang · S. Rangel · P. G. Lemaux
Department of Plant and Microbial Biology,
University of California, Berkeley,
CA 94720, USA

A. K. Sturbaum · V. Edwards · P. Bregitzer
National Small Grains Germplasm Research Facility,
USDA-ARS, Aberdeen, ID 83210, USA

K. Johnson · A. Kleinhofs
Department of Crop and Soil Sciences,
School of Molecular Biosciences,
Washington State University, Pullman,
WA 99164-6420, USA

V. Carollo
Western Regional Research Center, USDA-ARS, Albany,
CA 94710, USA

Introduction

Barley is an economically important crop and is ideal for implementing a reverse genetics gene-tagging system in the Triticeae. One of the main advantages of barley for genetic studies is that it is a true diploid (2n = 14) with seven cytologically distinct chromosomes containing approximately 5.3×10^9 bp of DNA (Bennett and Smith 1976). As such, recessive mutations can be more easily identified in barley than in its polyploid relatives, wheat and oats (Scholz et al. 2001). Barley chromosomes are homeologous with those of wheat (Moore et al. 1995), which allows barley to serve as a model system for the polyploid members of the Triticeae. In addition, barley has no known endogenous transposable elements (Scholz et al. 2001), although there are large numbers of

retrotransposon elements (Waugh et al. 1997; Kalendar et al. 2000).

Although the large size of the barley genome has precluded whole genome sequencing to date, the barley research community has amassed an array of other genomics resources, recently summarized by Hayes et al (2003), including genetic stocks, linkage maps, quantitative trait loci (QTL) databases, expressed sequence tags (ESTs), Single Nucleotide Polymorphisms (SNPs) (Kota et al. 2003), the Affymetrix barley microarray (Wanamaker and Close 2003) and a large-insert Bacterial Artificial Chromosome (BAC) library (Yu et al. 2000). More than 30 well-characterized genetic linkage maps are available (http://wheat.pw.usda.gov/ggpages/map_summary.html). These molecular marker linkage maps have been useful for identifying QTLs (Hayes et al. 2003) and will be essential for mapping the more than 360,000 ESTs, most of which still have no assigned function (Wanamaker and Close 2003). The barley BAC library covers 6.3 genome equivalents (Yu et al. 2000) and will ultimately allow for integration of the physical and genetic maps.

Maize *Activator* (*Ac*) and *Dissociation* (*Ds*) elements have been utilized as mutagens in heterologous systems since the cloning of the transposons themselves (Fedoroff et al. 1983) and the demonstration that these elements retain the ability to transpose in heterologous transgenic plants (Baker et al. 1986). Two-element (*Ac*/*Ds*) systems have several advantages over those based upon the use of *Ac* alone. In a two-element system the transposase gene (*Ac*TPase) can be segregated away from the non-autonomous *Ds* element, thereby stabilizing the non-autonomous element after its transposition to a new location (Koprek et al. 2000). Based upon the ability of *Ds* to transpose to linked sites (Parinov et al. 1999; Ito et al. 2002), targeted mutagenesis of loci genetically linked to one of the mapped *Ds* elements can be accomplished by reactivating the *Ds* element in crosses with transposase expressing lines. This may allow researchers to clone genes that would be difficult to access by map-based cloning, if the *Ds* element maps close to the gene of interest. In addition, initial *Ds* transpositions that land in genes of interest can be utilized directly, facilitating the functional analysis and cloning of those genes.

Significant steps have been taken towards the development of a transposon tagging system in barley. McElroy et al. (1997) demonstrated excision of the maize transposon *Ds* in barley scutellar tissues, using a simple transient assay in which *Ac*TPase was introduced simultaneously with a *uidA* (*β*-glucuronidase) gene disrupted by *Ds*. The development of an efficient transformation system for barley, using the cultivars Igri (Jähne et al. 1994) and Golden Promise (Wan and Lemaux 1994), has made it possible to employ this system to tag genes in barley. Koprek et al. (2000, 2001) have provided evidence of high transposition frequencies of *Ds* elements in a set of *Ds* -containing plants, after crosses with other plants expressing *Ac*TPase. Recently, Scholz and co-workers (2001) also demonstrated successful transposition of *Ac* in barley, with the ultimate goal of developing a one-element transposon tagging system.

The set of *Ds* insertions described in this study have been constructed in the variety. Golden Promise, in which successful transformation of barley was first accomplished (Wan and Lemaux1994), but Golden Promise is not a parent of any of the well-characterized mapping populations that are currently available (http://wheat.pw.usda.gov/ggpages/map_summary.html).
Thus, we are utilizing the Oregon Wolfe Barleys (OWBs), a population of 94 doubled haploids derived from the F_1 of a cross between the Dominant (OWB-D) and Recessive (OWB-R) Wolfe Multiple Marker Stocks (Costa et al. 2001). Integrating the virtual bin map Restriction Fragment Length Polymorphism (RFLP) markers of Kleinhofs and Graner (2001) into the OWB map will enable the OWBs to be used to connect the mapped *Ds* elements to the array of available genomic resources in barley. In the virtual bin map, each chromosome (linkage group) is divided in approximately 10 cM bins delineated by specific RFLP markers that have been mapped in multiple populations. With the addition of this core set of RFLP markers that define the bins, as described in this paper, the OWB population can serve as a central reference for integrating information from various barley genetic maps and allow the relative positions of mapped *Ds* insertion elements to be determined in other populations.

In addition to bin map integration, the OWBs offer other advantages for mapping the *Ds* insertions. The OWB doubled haploid (DH) mapping population is an immortal resource, as the lines are homozygous and the population is highly polymorphic (Costa et al. 2001). The OWB linkage map contains multiple types of markers, including morphological, RFLP, Simple Sequence Repeat (SSR) and Amplified Fragment Length Polymorphism (AFLP), as well as a sequence tagged site (STS) marker (*hvknox3*). The OWB DH population is also the focus of EST mapping efforts (Thiel et al. 2003). Additional information on the OWB population, which is an internationally utilized resource for barley research and instruction, is available at http://barleyworld.org and on GrainGenes http://wheat.pw.usda.gov/ggpages/map_summary.html.

We utilized a sequence-based strategy for assigning linkage map coordinates on the OWB map to each *Ds* insertion in Golden Promise. The principal strategy involved sequencing the genomic DNA flanking the *Ds* insertions in each Golden Promise TNP line, then cloning and sequencing the corresponding alleles from OWB-D and OWB-R (or other barley varieties, as necessary). We hypothesized that alignment of the OWB-D and OWB-R sequences would reveal SNPs or insertion/deletion (indel) polymorphisms that could be scored on the OWB mapping population. By locating the *Ds* insertions on the OWB map, and with reference to the bin map, researchers interested in a specific region of the

genome will be able to choose the appropriate Golden Promise *Ds* insertion line for their research. The TNP lines will be available upon request for non-commercial purposes through a Material Transfer Agreement (MTA) by accessing the Barley TNP link at Graingenes (http://wheat.pw.usda.gov/barleyTNP/).

Our long-term goal is to develop and make available to the research community a set of *Ds* insertion lines in barley, each containing a single, mapped, *Ds* element assigned to a linkage map bin. Our major objective in this phase of the project was to determine the locations of the *Ds* insertions in a number of Golden Promise TNP lines, utilizing the OWB population, to provide a preliminary characterization of those insertion sites using BLAST searches, and to address the mapped flanking sequences to BAC clones. Supporting objectives were the addition of anchor RFLP loci to the OWB map to facilitate integration with bin map resources, and assessment of the effectiveness of the sequence-based mapping strategy. In this paper we report the genetic mapping and preliminary characterization of 19 *Ds* insertions in the variety Golden Promise. In addition, we demonstrate that the OWB population can serve as a central reference map to integrate genomics resources available in the barley research community and that the sequence-based mapping strategy is effective.

Materials and methods

Development of the TNP lines in Golden Promise

Stable, single-copy *Ds* insertion lines were previously developed using a vector containing the phosphinothricin acetyl transferase gene (*bar*), under the control of the maize *ubiquitin* promoter and the *nos* terminator and flanked by the 5′ and 3′ regions of the *Ds* element that includes the inverted repeats (Koprek et al. 2000, 2001). Initially, *Ds* (pSP- *Ds* -Ubi-bar) and *AcTPase* (pBS-codA- *Act* -UbiAc) constructs were co-transformed into the barley cultivar Golden Promise by particle bombardment. Similarly, individual lines, which only contained and expressed *Ac*TPase, were also generated (Koprek et al. 2000). Based on DNA hybridizations, lines were identified that contained single (32B) or low-copy-number (32A/B) *Ds* elements, and these were crossed with Golden Promise lines expressing transposase. A number of new lines from these crosses were identified which contained transposed, single-copy, *Ds* insertions. Transposition was detected using a PCR assay to test for the presence of an empty donor site originating from the original plasmid construct, in combination with Southern analysis as described previously (Koprek et al. 2001). In this report, we use the nomenclature of Koprek et al. (2001) to describe the *Ds* insertion lines. Plants containing a single copy of transposed *Ds-bar* that had segregated away from the gene encoding *Ac*TPase, and other copies of *Ds-bar,* were designated "TNP" (transposed). The genomic integration sites of individual transposed *Ds* elements were given a " *DsT* " locus designation, followed by a number corresponding to the number of the TNP line.

DNA isolation

Genomic DNA was isolated from TNP lines as described by Dellaporta (1994) (for inverse PCR), and from OWB-D, OWB-R, Golden Promise, Dicktoo, Morex, Steptoe, the Dicktoo x Morex doubled haploid lines and TNP lines (for TAIL PCR) using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA). DNA concentrations were measured by spectrophotometry and by comparison with standard lambda DNA after electrophoresis in 1% agarose gels. For the OWB DH mapping population, extraction and purification of DNA were performed using a modification of the method of Liu and Whittier (1994). Samples of young leaves or hypocotyl tissue (30–50 mg) were ground in 400 µl of cold nuclear isolation buffer, with the addition of 2% (w/v) Sarkosyl, using a mixer mill (Retsch MM 300; Retsch, Newtown, PA). Samples were extracted with phenol: chloroform and precipitated with salt and ethanol (Sambrook and Russell 2001). The pelleted DNA was resuspended in 200 µl of TE (pH 8.0) containing RNase (10 µg/ml) and incubated overnight at 4°C. DNA concentrations were determined as above.

Nested inverse PCR

Inverse PCR (iPCR) (Ochman et al. 1988, 1993) was performed as described by Li et al. (2001), with the following modifications. Genomic DNA (5–10 µg) was digested either with *Nco* I or *Nhe* I, followed by heat-inactivation of the enzyme, extraction with phenol:chloroform (1:1), and self-ligation. The iPCRs contained approximately 100 ng of purified, ligated DNA, 0.2 mM dNTP mixture, each primer at 0.1 µM and 2.5 U of *ExTaq* DNA polymerase (TaKaRa Bio, Shiga, Japan) in a 50-µl reaction mixture. The iPCR mixture was heated at 94°C for 5 min before adding the *ExTaq* DNA polymerase. Thirty six cycles of PCR were carried out, with each cycle consisting of 94°C for 45 s, 60°C for 45 s and 72°C for 150 s. In the first round of PCR designed to obtain flanking sequences, primers P3 (5′-GTAGATAATGCCAGCCTGTT-3′) and either P1 (5′-ATGTGCTACATTAAACTATG-3′) or P2 (5′-CGACCGGATCGTATCGGT-3′) were used on the 5′ side, and the primers P6 (5′-CATATTGCAGTCATCCCGAA-3′) and either P4 (5′-TGCGGAACGGCTAGAGCCAT-3′) or P5 (5′-ACAGGTCGCATCCGTGTACGAACG-3′) were used on the 3′ side of the *Ds* element. The PCR products from the first round of PCR were diluted (20–50-fold) with purified water and used as the template for the second round.

Specific PCR products were generated in the second round of PCR by amplifying with the nested primers P8 (5′-CTCGTGTTGTTCGGAGCGCACACA-3′) and P7 (5′-TTCGTTTCCGTCCCGCAAGT-3′) on the 5′ side, and P11 (5′-AACTAGCTCTACCGTTTCCG-3′) and P5 (see above) or P10 (5′-TAGCAGCACGGATCTAACAC-3′) on the 3′ side. The iPCR products were gel-purified using the QIAquick gel extraction kit (Qiagen) and sequenced commercially by Elim Biopharmaceuticals (Hayward, CA).

TAIL PCR

Thermal asymmetric interlaced (TAIL)-PCR was performed essentially as in Liu et al. (1995) and Koprek et al. (2000) with the following modifications. The degenerate primers were AD1 (Liu et al. 1995), AD1 (as in Koprek et al. 2000), AD2, AD3 and AD6, all from Koprek et al. (2000), and the nested *Ds* -specific primers were *Ds* A, *Ds* B and *Ds* C from Koprek et al. (2000). For each reaction, 25 ng of genomic DNA was used with the first *Ds* primer at a final concentration of 1 μM in combination with the respective TAIL PCR primer. After the primary amplification, PCR products were diluted 50-fold and used as the template for the secondary reaction (500-fold final dilution). After the secondary amplification, PCR products were diluted 10-fold and used as templates for the tertiary reaction (100-fold final dilution). For the tertiary reaction, we repeated the TAIL 2 PCR program (Liu et al. 1995). The annealing temperatures used in the TAIL PCR programs were modified from those used by Liu et al. (1995) to match the melting temperatures specific to the nested *Ds* -specific primers. After the tertiary amplification, PCR products were gel-purified, as described above. The gel-purified amplification product was cloned into the TOPO TA cloning vector, pCR2.1-TOPO (Invitrogen, Carlsbad, CA) and sequenced at the Center for Gene Research and Biotechnology (Oregon State University, Corvallis, OR). Sequences were analyzed using BioEdit Sequence Alignment Editor (Hall 1999).

Mapping *DsT* loci using sequence polymorphisms

Primers, based on the sequences flanking *Ds* elements in the TNP lines, were designed to identify polymorphisms in the sequences of the corresponding alleles from OWB-D and OWB-R (Table 1). Amplification products were also generated from wild-type Golden Promise and the parents of other mapping populations (Dicktoo, Morex and Steptoe), if needed. Wherever possible, primers were designed in such a way that the amplicons would span the *Ds* insertion site, but in some cases, flanking sequence was only obtained from either the 5′ or 3′ side, relative to the *Ds* element. Amplification products were generated using Platinum *Taq* DNA polymerase High Fidelity or Platinum *Taq* DNA polymerase (Invitrogen). At least two clones from each genotype were sequenced in each direction. With clones that were less than 200 bp in length, sequencing was performed in only one direction.

Sequences were compared in BioEdit 4.8 (Hall 1999) to identify SNPs or indels that were consistent between the sequenced clones of each genotype. Where a SNP was identified, a Cleaved Amplified Polymorphic Sequence (CAPS) assay was designed to differentiate between the parental alleles. The corresponding sequence was amplified from each member of the OWB DH mapping population, and digested with the appropriate restriction enzyme. Indels were assayed after amplifying the corresponding sequences from each member of the mapping population. Depending upon the size of the DNA fragments, the individual members of the appropriate mapping population were assayed and scored for the polymorphisms on 2–3% agarose gels stained with ethidium bromide.

RFLP assays

Eight barley genotypes (OWB-D, OWB-R, Steptoe, Morex, Harrington, Q-line, Foster and C14196) that are the parents of five mapping populations (Steptoe × Morex, Harrington × Morex, Steptoe × Q-line, OWB-D × OWB-R and Foster × C14196) were each digested with six restriction enzymes (*Bam* HI, *Dra* I, *Eco* RI, *Eco* RV, *Hind* III, *Xba* I) and hybridized to probes generated using the cloned genomic DNA flanking the *Ds* insertions in TNP-11, 30, 32B and 41 (Tables 1 and 2). The procedures used for DNA extractions and RFLP analyses are described in Kleinhofs et al. (1993).

Linkage map construction and alignment with the bin map

The most recent version of the OWB map (http://www.barleyworld.org, posting date July 24, 2002) that contained 152 markers [78 RFLPs, 61 SSRs, 11 morphological markers with one STS marker (*hvknox3*) and one mildew (*Blumeria graminis*) resistance gene (*RbgMD*)] was used as the starting point for constructing an updated linkage map. Forty seven newly-scored RFLP loci (A. Kleinhofs, Washington State University, Pullman, Wash.) were added to align the OWB map with the virtual bin map, together with three additional morphological markers. Hairy rachis (*Hrq7*) and glume hair (*GlhMD*) were scored by visual assessment (K. Sato, Barley Germplasm Center, Research Institute for Bioresources, Okayama University, Kurashiki 710, Japan), while rough awn (*raw1*) was scored by tactile assessment of awns from tip to base and confirmed by examining awns under a dissecting microscope. Three additional STS markers: a putative cytochrome P450 gene

Table 1 List of primers used to generate sequences for mapping Ds insertions and to generate probes for RFLP analysis and screening the Morex BAC Library

Locus	Primer pairs sequences (5' to 3')	Source of template	Accession No. of product
DsT-1	TNP-1-91F/TNP-1-1392R	OWB-D	AY505351
	(AGAATGAGGAAAGGGAATGG)	OWB-R	AY505352
	(AGTTACCGGACTCGGACTGG)		
DsT-1	TNP-1-85F/TNP-1-1392R	Golden Promise	AY505353
	(TCGGATAAACGAACATACAGC)		
	(AGTTACCGGACTCGGACTGG)		
DsT-6	TNP-6-495F/TNP-6-922R	OWB-D	AY505357
	(AACATACACCTCTGTTCCATCATGC)	OWB-R	AY505358
	(TTTTTTGCCGGACTTTTTCTATGTCG)	Golden Promise	AY520812
DsT-11	TNP-11-13F/TNP-11-515R	OWB-D	AY505359
	(TTTTGCAGATGCCACTTGAG)	OWB-R	AY505360
	(CAACTTCTCAACGACGGACA)		
DsT-11	TNP-11-69F/TNP-11-219R	Golden Promise	AY505361
	(TGCGCGTATCGGTGTATGTA)	Dicktoo	AY505362
	(CCCTCCGTCCCAAAATAAAT)	Morex	AY505363
DsT-13	TNP-13-19F/TNP-13-637R	OWB-D	AY505366
	(GTTCACCTGCTCCATAACATGCG)	OWB-R	AY505367
	(ACTGGGTGCCCCACTGTTAGG)	Golden Promise	AY505368
DsT-18	TNP-18-168F/TNP-18-802R	OWB-D	AY505375
	(CCGGTCCATGGTGTTGGCAAG)	OWB-R	AY505376
	(GACAGCCCCCAAAACAGTTCAATG)	Golden Promise	AY505377
DsT-22	TNP-22-5F/TNP-22-385R	OWB-D	AY505378
	(AGAGGGAAAGCTCCTCTTGG)	OWB-R	AY505379
	(TGTTGATGAGACCAGCCAAT)	Golden Promise	AY505380
DsT-24	TNP-24-751F/TNP-24-1250R	OWB-D	AY505381
	(GTGGTAGTGCTGGCAGTTCA)	OWB-R	AY505382
	(TGGAGAACAAAGGATAGGCTGT)	Golden Promise	AY594686
DsT-27	TNP-27-42F/TNP-27-627R	OWB-D	AY505383
	(GTACAGGCCAGCCACAAGAACATG)	OWB-R	AY505384
	(CCATAACCCAACCAAAGACCACC)	Golden Promise	AY505385
		Dicktoo	AY505386
		Morex	AY505387
		Steptoe	AY505388
DsT-28	TNP-28-66F/TNP-28-517R	OWB-D	AY505389
	(TTTTGGAGGGATTGAAGGTG)	OWB-R	AY505390
	(AGTCTATAGGCGTCTCTTGTGG)	Golden Promise	AY505391
DsT-29	TNP-29-9F/TNP-29-785R	OWB-D	AY505413
	(GGAGAGCCACCCTTATGTGA)	OWB-R	AY505414
	(TCGAAAACCAAGTGCAAACC)	Golden Promise	AY505415
DsT-30	TNP-30-24F/TNP-30-403R	OWB-D	AY505392
	(AAAACGGGCAGTCCAAGTGC)	OWB-R	AY505393
	(TGTTCTCTGCCAGCAACTTATGC)	Golden Promise	AY505394
DsT-31	TNP-31-133F/TNP-31-748R	OWB-D	AY505395
	(GGACATCGCATCGCCTACAC)	OWB-R	AY505396
	(CACACCCTTTCAACCCCCAAG)	Golden Promise	AY505397
DsT-32B	TNP-32B-39F/TNP-32B-713R	OWB-D	AY505398
	(TGTGCAGGTGGTTTCCAAG)	OWB-R	AY505399
	(GCCACCGACCTTCTTGTTGG)	Golden Promise	AY505400
DsT-33	TNP-33-24F/TNP-33-348R	OWB-D	AY594683
	(GCACACATATTATCATGAAAAAGAGC)	OWB-R	AY594684
	(ACCCCAAATGAGTTTCGATG)	Golden Promise	AY594685
DsT-34	TNP-34-119F/TNP-34-533R	OWB-D	AY505401
	(CCTCTACCAAAATGGTGCTCTTGACG)	OWB-R	AY505402
	(CCAGTCCGGTGAATTTCAGGGATAC)	Golden Promise	AY505403
		Dicktoo	AY505404
		Morex	AY505405
		Steptoe	AY505406
DsT-35	TNP-35-136F/TNP-35-658R	OWB-D	AY505407
	(CAATAGTGGGTGGGGCAAGC)	OWB-R	AY505408
	(CAGTGCAACTTGGCGAGACG)	Golden Promise	AY505409
DsT-40	TNP-40-56F/TNP-40-349R	OWB-D	AY629352
	(CGAGACCGACGGAAGTAGG)	OWB-R	AY629353
	(CAATAGAATTGGTGGGGAAGC)	Golden Promise	AY629354
DsT-41	TNP-41-1006F/TNP-41-1469R	OWB-D	AY629355
	(GGTGGCCGGTGTTACCAG)	OWB-R	AY629356
	(ACACACCACTTATCGCAACC)	Golden Promise	AY629357
DsT-53	TNP-53-106F/TNP-53-464R	OWB-D	AY629358
	(CGGTCCTCAGAGTCATCTCG)	OWB-R	AY629359
	(CCCATGAGCACGTATTGTCC)	Golden Promise	AY629360

aOregon Wolfe Barley Dominant Multiple Marker stock
bOregon Wolfe Barley Recessive Multiple Marker stock

(*OSU-STS1*), and two uncharacterized sequences (*OSU-STS2 and OSU-STS3*) were mapped in the OWB population (L. Marquez-Cedillo and L. Cooper, Oregon State University, Corvallis, Ore.).

All previously mapped and newly scored markers were assessed for missing data, cosegregation and segregation distortion. Markers with more than 15% missing data or showing significant segregation distortion ($P < 0.05$) were discarded and were added back only if they were essential for maintaining a united linkage group (as in the case of ABG004 on chromosome 3H). In addition, RFLP markers that did not define bins, but cosegregated with an SSR or a *DsT* locus, were discarded. The RFLPs that defined bins were retained, as were all morphological markers. Linkage map construction was performed using JoinMap 3.0 (Van Ooijen and Voorrips 2001), adjusting LOD and recombination settings as necessary to maintain the integrity of the linkage groups.

Addressing *DsT* loci to BAC clones

Nineteen genomic clones (Table 2) corresponding to mapped *DsT* loci were assigned to BAC clones by hybridization to a subset of four of the 17 filters (except for TNP-11, 27, 32B, 33, 40 and 53 which were each hybridized to 8 of the filters) comprising the Morex genomic BAC library (Yu et al. 2000). Probes were produced by PCR amplification from individual clones in pCR 2.1-TOPO (Invitrogen) using primers specific to the genomic DNA flanking the *Ds* insertions. PCR products were purified by agarose gel extraction to avoid contamination with plasmid DNA. Probe labeling was performed by random priming (Feinberg and Vogelstein 1983, 1984) using the Megaprime DNA labeling system (Amersham Biosciences, Piscataway, NJ), following the manufacturer's recommendations. Hybridizations to the BAC filters were performed as recommended by the filter supplier (http://www.genome.clemson.edu/groups/

Table 2 Summary of *Ds* insertion site mapping in barley based on chromosome and bin location, mapping strategy, number of positive BAC clones identified, cereal EST matches and E values

DsT insertion	Chromosome bin[a]	Mapping strategy[b]	Number of positive BAC clones[c]	Cereal EST matches[d]	E value[e]
1	2H-1	Indel OWB-R (45 bp insertion)	16	Maize EST CB885436	2.3e-7
6	5H-1	indel OWB-R (31 bp deletion)	5	Wheat EST BE443030	6.0e-16
11	5H-10	Indel Dicktoo (16 bp deletion), verified by RFLP Steptoe x Q-Line Mini (*Hind*III, single copy)	32	Barley EST AL505601 (putative wall-associated protein kinase)	2.9e-141
13	5H-14	SNP OWB-R (*Acil* I)	4	Barley EST BF266272.2	6.5e-5
18	6H-12	indel OWB-D (228 bp deletion)	2	Wheat EST Contig TC94268	4.8e-28
22	6H-13	SNP OWB-D (*Rsa* I)	6	Barley Array Contig No. 10453; similar to putative aristolochene synthase (*Oryza sativa*)	1e-111
24	3H-14	SNP OWB-D (*Mnl* I)	4	Rice EST Contig TC128765	6.1e-89
27	3H-4	SNP OWB-D (*Aci* I)	1	Unknown	3.3e-2*
28	6H-14	SNP OWB-R (*Aci* I)	1	Unknown	2e-3*
29	4H-3	SNP OWB-R (*Mse* I)	1	Barley Array Contig NO. 2257; similar to putative EREBP-type transcription factor (*Oryza sativa*)	4e-8
30	7H-9	SNP OWB-D (*Taq* I), verified by RFLPOWB-D x OWB-R (*Xba*I, single copy)	6	Unknown	0.13*
31	7H-3	SNP OWB-R (*Cac* 8I)	3	Unknown	0.35*
32B	6H-12	SNP OWB-D (*Pst*I), verified by RFLPOWB-D x OWB-R (*Hind*III, 2–3 copies)	3	Wheat EST AL818129	2.5e-10
33	5H-14	indel OWB-D (20-bp deletion)	11	Unknown	2.0e-3*
34	5H-2	SNP Steptoe (*Sac* II)	1	Unknown	9.2e-2*
35	3H-16	SNP OWB-R (*Bsp1286*I)	2	Barley Array Contig No. 39284	5e-4
40	3H-15	SNP OWB-D (*Hinc* II)	3	Barley Array Contig No. 14757	9.4e-2*
41	2H-10	SNP-OWB-R (*Mnl* I), verified by RFLPOWB-D x OWB-R (*Xba* I, single copy)	2	Rice TC196496 (probable ubiquitin conjugating enzyme)	2.1e-42
53	7H-2	SNP OWB-R (*Pflm* I)	4	Wheat alpha-gliadin storage protein gene	4e-08

[a] *DsT* loci were assigned bin locations based upon the OWB bin map (Fig. 1)

[b] See text for full details of the mapping strategies

[c] Genomic clones corresponding to mapped *DsT* loci were hybridized to a subset of the 17 filters comprising the Morex genomic BAC library (Yu et al. 2000). The number of positive hybridizations is shown. See the Electronic Supplementary Material for the BAC addresses of the clones that hybridize to each *DsT* probe

[d] BLASTN and TBLASTX searches were performed using genomic DNA sequences from wild type Golden Promise corresponding to the DNA flanking each *Ds* insertion event

[e] For the purposes of this Table, the best match from all the searches is shown, although any match with an E value greater than $1e{-}4$ is not considered to be significant, so the putative identities of those are not shown. *Asterisks* indicate E values below the significance threshold

bac/protocols/addressnew.html) by adapting the hybridization temperatures to the melting temperature of each cloned fragment.

Characterization of insertion site sequences

BLASTN and TBLASTX (Altschul et al. 1990, 1997) searches were performed using genomic DNA sequences from wild-type Golden Promise corresponding to the DNA flanking each *Ds* insertion event. The following databases were utilized: The Institute for Genomic Research (TIGR) gene indicies for barley, *Arabidopsis*, maize, rice and wheat (http://tigrblast.tigr.org/tgi/), the EST sequences used for the Affymetrix barley microarray HarvEST 1.12 (Wanamaker and Close 2003) (http://138.23.191.152:/blast/blast.html) and the non-redundant NCBI database on GenBank (http://www.ncbi.nlm.nih.gov/BLAST/). The most significant Expectation (E) values from the various searches are presented in Table 2.

Accession numbers

The sequences used for mapping (see also Table 1) are available from GenBank under the Accession Nos.: AY505351–AY505353, AY505357–AY505363, AY505366–AY505371, AY505375–AY505409, AY505413–AY505415, AY520812 and AY594683–AY594686, AY629352–AY629360.

Results and discussion

Development of the TNP lines in Golden Promise

A single-copy *Ds* line in Golden Promise, TNP-32B, was identified and was crossed with Golden Promise lines expressing transposase (Koprek et al. 2001). The location (chromosome 6H, bin 12) of the original *DsT* insertion was determined by mapping an SNP in the genomic DNA flanking the insertion site (see Table 2), which was verified independently by RFLP analysis. From these crosses, the lines TNP-6–TNP-30 and TNP-40, 41 and 53 were identified, all of which contained transposed, single-copy, *Ds* insertions. Transposition from the original locus was detected using a PCR assay for the presence of an empty donor site, and confirmed by Southern hybridization analysis (Koprek et al. 2001).

Among the lines generated by the reactivation of the *Ds* insertion in TNP-32B, three (*DsT-18*, *22* and *28*) carry the transposed insert at positions which are tightly linked to the site of the original *Ds* insertion, *DsT*-32B, on chromosome 6H. Other transpositions are located on different chromosomes, with one each on chromosomes 2H (*DsT-41*) and 4H (*DsT-29*), two on 7H (*DsT-30* and *53*) and three each on chromosomes 3H (*DsT-24*, *27* and *40*) and 5H (*DsT-6*, *11* and *13)*. Although the numbers of *Ds* insertions surveyed here are too small to allow us to make generalizations, it appears as though the *Ds* elements have transposed to both linked and unlinked positions in the barley genome. These results support the hypothesis that, in barley, as observed in other heterologous plant systems such as *Arabidopsis* (Parinov et al. 1999; Ito et al. 2002) and rice (Kolesnik et al. 2004), the *Ds* element transposes into both linked and unlinked loci.

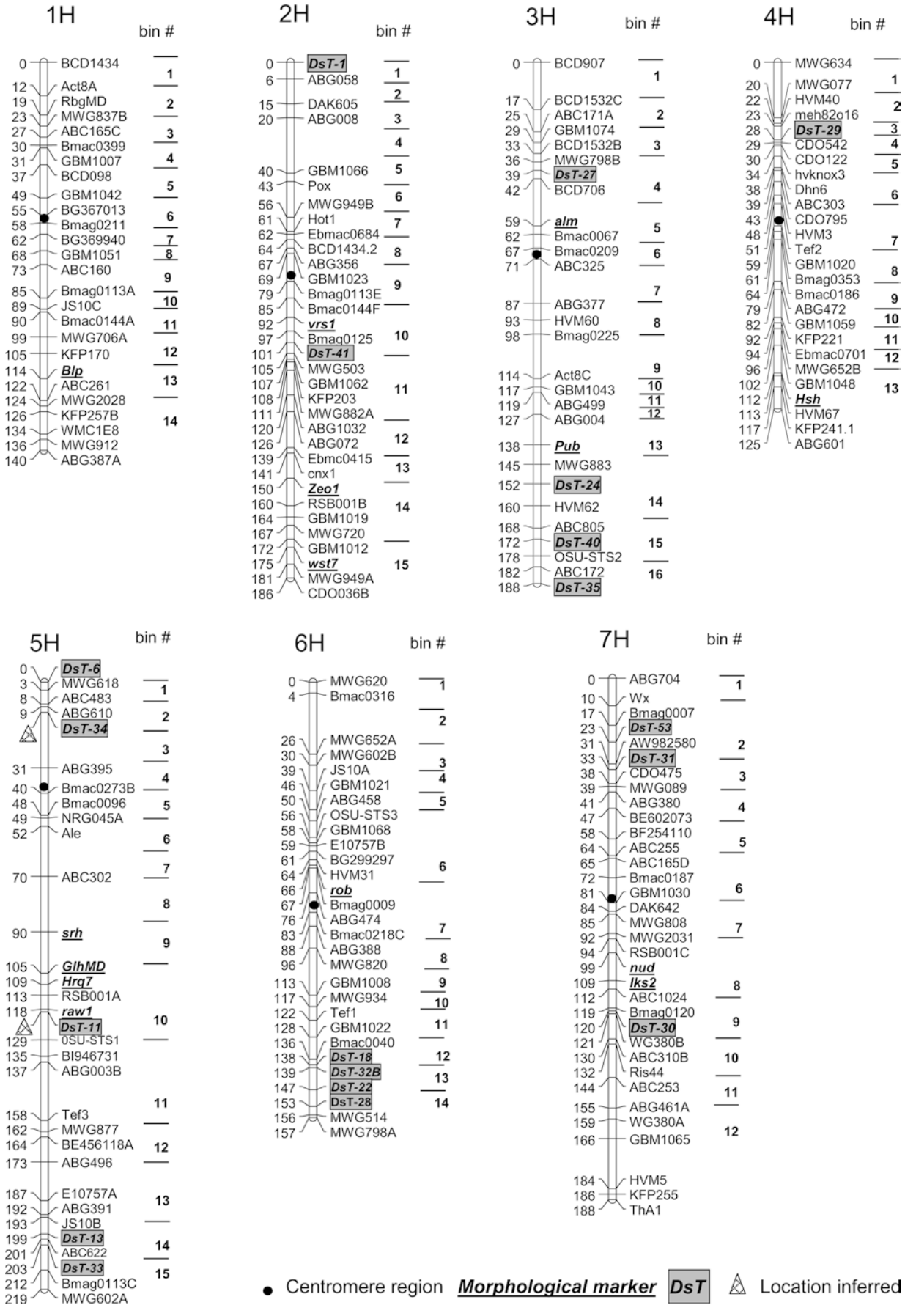
The original parent for a smaller set of transposition lines (TNP-1, 31 and 33–35) was TNP-32A/B, a multicopy *Ds* parent. The *Ds* insertions in this line were reactivated as above, but the genetic map location of the progenitor *Ds* insertions in TNP-32A/B could not be determined. Single-copy transpositions from this set of crosses were mapped to chromosomes 2H (*DsT-1*), 3H (*DsT-35*), 5H (*DsT-33*) and 7H (*DsT-53*), respectively.

We were interested to see if any TNP lines had acquired a visible mutant phenotypes as a result of the insertion of the *Ds*. No obvious, visible changes in phenotypes were observed, although we predict that some of the *Ds* insertions may result in changes in biochemical or proteomic profiles and new patterns of gene expression.

Summary of mapping results and integration with the bin map

We assigned linkage map positions to *Ds* insertions in 19 Golden Promise TNP lines by mapping polymorphisms in the genomic DNA flanking each *Ds* insertion. Seventeen *DsT* loci polymorphisms were scored in the OWB DH population and directly placed on the OWB linkage map (Fig. 1). The positions of remaining two *DsT* loci were inferred based on bin assignments. In these cases, we did not identify any polymorphisms in the OWB-D or OWB-R DNA corresponding to the flanking sequence, so we utilized other mapping populations. *DsT-34* was mapped in the Steptoe × Morex mini population (Mgonja et al. 1995), while *DsT-11* was mapped in the Dicktoo × Morex population (Pan et al. 1994). The inferred linkage map positions of these loci were assigned based on the presence of markers common to the OWB population in the bin map.

The OWB map (Fig. 1) now includes 206 markers: 14 morphological markers, 52 SSRs, 118 RFLPs, one mildew (*Blumeria graminis*) resistance gene (*RbgMD*), four STS markers (*hvknox3* and *OSU-STS1–3*) and 17 *DsT* loci (with two additional *DsT* loci in inferred positions). The order of SSRs agrees with the maps of Ramsay et al. (2000) and Thiel et al. (2003), and the order of RFLP markers corresponds to that in the virtual bin map of Kleinhofs and Graner (2001). Fifty-seven RFLPs were useful in assigning bins directly to the OWB map (Fig. 1). The remaining RFLPs were useful because their bin locations can be inferred based on mapping


1H
bin #
0 BCD1434
12 Act8A
19 RbgMD
23 MWG837B
27 ABC165C
30 Bmac0399
31 GBM1007
37 BCD098
49 GBM1042
55 BG367013
58 Bmag0211
62 BG369940
68 GBM1051
73 ABC160
85 Bmag0113A
89 JS10C
90 Bmac0144A
99 MWG706A
105 KFP170
114 Blp
122 ABC261
124 MWG2028
126 KFP257B
134 WMC1E8
136 MWG912
140 ABG387A
2H
bin #
0 DsT-1
6 ABG058
15 DAK605
20 ABG008
40 GBM1066
43 Pox
56 MWG949B
61 Hot1
62 Ebmac0684
64 BCD1434.2
67 ABG356
69 GBM1023
79 Bmag0113E
85 Bmac0144F
92 vrs1
97 Bmag0125
101 DsT-41
105 MWG503
107 GBM1062
108 KFP203
111 MWG882A
120 ABG1032
126 ABG072
139 Ebmc0415
141 cnx1
150 Zeo1
160 RSB001B
164 GBM1019
167 MWG720
172 GBM1012
175 wst7
181 MWG949A
186 CDO036B
3H
bin #
0 BCD907
17 BCD1532C
25 ABC171A
29 GBM1074
33 BCD1532B
36 MWG798B
39 DsT-27
42 BCD706
59 alm
62 Bmac0067
67 Bmac0209
71 ABC325
87 ABG377
93 HVM60
98 Bmag0225
114 Act8C
117 GBM1043
119 ABG499
127 ABG004
138 Pub
145 MWG883
152 DsT-24
160 HVM62
168 ABC805
172 DsT-40
178 OSU-STS2
182 ABC172
188 DsT-35
4H
bin #
0 MWG634
20 MWG077
22 HVM40
23 meh82o16
28 DsT-29
29 CDO542
30 CDO122
34 hvknox3
38 Dhn6
39 ABC303
43 CDO795
48 HVM3
51 Tef2
59 GBM1020
61 Bmag0353
64 Bmac0186
79 ABG472
82 GBM1059
92 KFP221
94 Ebmac0701
96 MWG652B
102 GBM1048
112 Hsh
113 HVM67
117 KFP241.1
125 ABG601
5H
bin #
0 DsT-6
3 MWG618
8 ABC483
9 ABG610
DsT-34
31 ABG395
40 Bmac0273B
48 Bmac0096
49 NRG045A
52 Ale
70 ABC302
90 srh
105 GlhMD
109 Hrq7
113 RSB001A
118 raw1
DsT-11
129 OSU-STS1
135 BI946731
137 ABG003B
158 Tef3
162 MWG877
164 BE456118A
173 ABG496
187 E10757A
192 ABG391
193 JS10B
199 DsT-13
201 ABC622
203 DsT-33
212 Bmag0113C
219 MWG602A
6H
bin #
0 MWG620
4 Bmac0316
26 MWG652A
30 MWG602B
39 JS10A
46 GBM1021
50 ABG458
56 OSU-STS3
58 GBM1068
59 E10757B
61 BG299297
64 HVM31
66 rob
67 Bmag0009
76 ABG474
83 Bmac0218C
88 ABG388
96 MWG820
113 GBM1008
117 MWG934
122 Tef1
128 GBM1022
136 Bmac0040
138 DsT-18
139 DsT-32B
147 DsT-22
153 DsT-28
156 MWG514
157 MWG798A
7H
bin #
0 ABG704
10 Wx
17 Bmag0007
23 DsT-53
31 AW982580
33 DsT-31
38 CDO475
39 MWG089
41 ABG380
47 BE602073
58 BF254110
64 ABC255
65 ABC165D
72 Bmac0187
81 GBM1030
84 DAK642
85 MWG808
92 MWG2031
94 RSB001C
99 nud
109 Iks2
112 ABC1024
119 Bmag0120
120 DsT-30
121 WG380B
130 ABC310B
132 Ris44
144 ABC253
155 ABG461A
159 WG380A
166 GBM1065
184 HVM5
186 KFP255
188 ThA1
Centromere region
Morphological marker
DsT
Location inferred

◄

Fig. 1 Oregon Wolfe Barley linkage and bin map. The map includes 19 mapped *Ds* insertions (*DsT*-), as well as morphological markers, RFLPs, SSRs and centromeric regions. Distances are given in Kosambi cM and are cumulative. *Dst-11* was mapped in the Dicktoo × Morex population; *DsT-34* in the Steptoe × Morex mini-population. The inferred positions of these loci are indicated by the *triangles* to the *left* of 5H

information from other populations. Ninety-three of the 99 bins defined by Kleinhofs and Graner (2001) can be directly aligned with the OWB map. The remaining six bins are internal to linkage groups and their positions can therefore be inferred. The complete coverage of the barley genome by the OWB map is critical for aligning the positions of the *DsT* loci onto maps from other cultivars and thus, for the subsequent selection of TNP lines for reactivation to clone genes of interest.

Mapping of the *Ds* insertions in the OWB population, together with consensus RFLP markers, places each *Ds* element in a specific bin. This will facilitate the integration of functional and structural genomics resources in barley. By assigning the *DsT* loci to bins on the OWB map, researchers can identify the *Ds* insertions closest to mapped morphological loci, genes of known function or QTL regions identified in other cultivars. These *Ds* insertions can then serve as targets for reactivation in order to tag linked genes.

The sequence-based mapping strategy also has the advantage of generating new sequence-based markers for the Triticeae research community. The sequences generated from OWB-D, OWB-R, wild-type Golden Promise, Dicktoo, Morex and Steptoe during mapping of *Ds* insertions have been deposited in GenBank. In some cases, the sequences submitted to GenBank are longer than those used for mapping, since shorter sequences that contain the polymorphisms are more accurately assayed. The PCR primers used to generate the sequences, and the associated GenBank Accession numbers, are provided in Table 1. To date, we have *DsT* loci in 17 of the 99 bins, on six of the seven chromosomes (Fig. 1, Table 2). Our long- term goal is to assign at least one *DsT* insertion to each bin.

Mapping of *DsT* insertions using CAPS to assay for SNPs

Fourteen *DsT* loci were mapped using a CAPS marker strategy to assay SNPs (*DsT 13, 22, 24, 27, 28, 29, 30, 31, 32B, 34, 35, 40, 41, 53*) using the appropriate restriction enzymes (Table 2). Of these, 13 (all except for *DsT-34*) were mapped directly in the OWB DH mapping population. SNPs are the most common type of sequence difference between alleles (Rafalski 2002) and have been used for mapping ESTs in maize (Batley et al. 2003) and barley (Kota et al. 2003). Although there are numerous technologies available for performing SNP assays in mapping populations (Rafalski 2002), we chose a restriction digest-based assay of PCR products (CAPS assay) as demonstrated by Konieczny and Ausubel (1993).

In order to illustrate the sequence-based mapping strategy, an example of the mapping procedures for a CAPS assay of an SNP (Fig. 2) is presented. Figure 2 illustrates a SNP in the 3′ flanking sequence of *DsT-30*. A 379-bp product from both the OWB-D and OWB-R genotypes was generated by PCR. Comparison of the sequences from OWB-D and OWB-R revealed three SNPs (at bases 196, 300 and 351). The SNP at postion 196 (T → C) creates a *Taq* I restriction site (T/CGA) in the OWB-D amplicon. *Taq* I digestion of OWB-D results in two products (184 bp and 195 bp, which co-migrate at the resolution afforded by a 2% agarose gel), in comparison to the 379-bp product in the OWB-R. This gives a clear and easily scored polymorphism, as shown for OWB-D, OWB-R and eight of the OWB DH lines (Fig. 2).

Mapping of *DsT* insertions using indels

Five of the *DsT* loci were mapped by assaying for insertions or deletions in the sequences flanking the *Ds* elements (Table 2) using agarose gel electrophoresis. *DsT-1, 6, 18* and *33* were assayed directly in the OWB DH mapping population, while the indel near *DsT-11* was mapped in the Dicktoo × Morex mapping population.

The mapping of *DsT-1* provides an example of the PCR-based detection of an indel (Fig. 3). The OWB-R

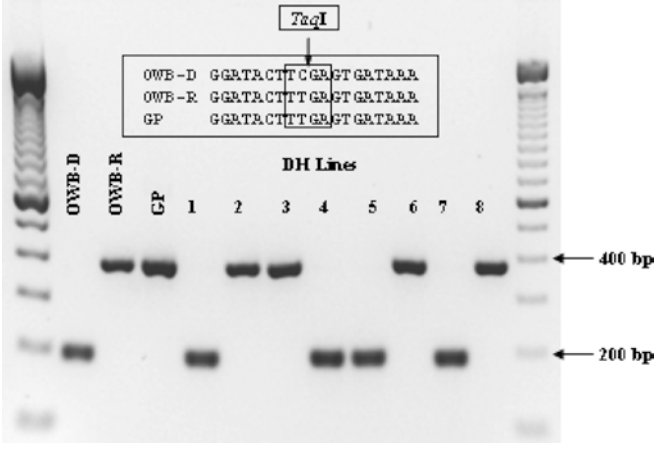


Fig. 2 A Single Nucleotide Polymorphism (SNP) in the TNP-30 OWB-D sequence creates a recognition site for *Taq* I. OWB-D, OWB-R and GP lines were amplified using TNP-30 24F/430R primers (see Table 1), producing a 379-bp amplicon in all three lines (a partial sequence alignment is shown). After cloning and sequencing, a SNP (T → C) was identified at position 196, which creates a *Taq* I restriction site (TC/GA) in the OWB-D amplicon. To assay the SNP in the OWB doubled haploid (DH) population, TNP-30 24F/430R amplification products from each DH line were digested with *Taq* I (DH lines 1–8 are shown here) and electrophoresed on a 2% agarose gel. *Taq* I digestion of DNA from OWB-D, as well as DH lines 1, 4, 5 and 7 resulted in complete digestion of the 379-bp amplification product

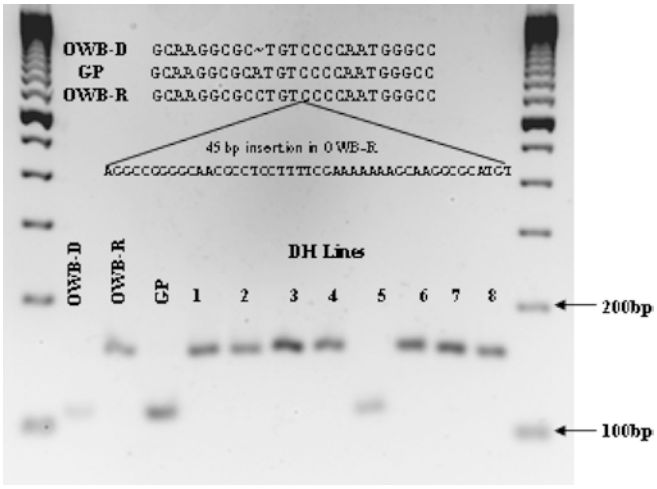


Fig. 3 A 45-bp insertion in the TNP-1 OWB-R sequence detected by PCR and agarose gel electrophoresis. DNAs from OWB-D, OWB-R and GP lines were amplified using TNP-1 425F/535R primers (see Table 1), producing a 110-bp amplicon from OWB-D and GP, as compared to a 155-bp product in OWB-R. After cloning and sequencing, a 45-bp insertion was identified in the OWB-R sequence (a partial sequence alignment is shown here). To assay for the presence of the insertion in the OWB DH population, TNP-1 425F/535R amplification products from each DH line (DH lines 1–8 shown here) were analyzed by electrophoresis on a 2% agarose gel. The 155-bp PCR product is evident in OWB-R, as well as in the DH lines 1–4 and 6–8

amplicon contained a 45-bp insertion relative to the OWB-D and Golden Promise wild-type sequences. The size polymorphism was scored in the OWB population after separating the amplification products on an agarose gel.

Confirmation of the sequence-based mapping results by RFLP analysis

DsT-11, *30, 32B* and *41* were mapped by RFLP analysis to confirm the results of the sequence-based mapping approach (Table 2). Where possible (*DsT-30, 32B* and *41*), the RFLP was mapped in the OWB population, and the RFLP data match the mapping data generated by the sequence-based approach, validating the results obtained by the CAPS-based assay of SNPs. In the case of *DsT-11*, the RFLP data cannot be directly compared to the data obtained by the sequence-based mapping approach, since the two assays were performed in different mapping populations, but the relative positions on the genetic linkage maps are the same. All four of the RFLP probes detected single- or low-copy-number sequences in the DNA hybridization analyses (Table 2). In addition to providing map locations, the results from the RFLP analysis indicated that all four of the *DsT* loci studied were located in non-repetitive, low-copy-number DNA sequences, as verified by the results of the BLAST searches of the sequence flanking the insertion site and by hybridization of the flanking sequence to BAC clones.

Comparison with other mapping strategies

Several other strategies could be used to map *Ds* insertions in the barley genome. Where the genomic sequence is known, as in the case of *Arabidopsis* (Parinov et al. 1999; Tissier et al. 1999; Ito et al. 2002; Raina et al. 2002) or rice (Kolesnik et al. 2004), it is relatively simple to assign *Ds* (or other transposable element) insertions to a position in the genome by comparison of genomic DNA sequences flanking the insertions with sequences of large insert clones of known physical or genetic position. Unfortunately, at the present time, the assignment of linkage map coordinates to barley BAC clones and ESTs is not sufficiently advanced to allow the use of this strategy in barley. Of the 313,344 BACs in the Morex library (Yu et al. 2000) few have been assigned a genetic map location, although a comprehensive effort to remedy this situation is now underway (T. Close, NSF Award No. 0321756).

The positions of *Ds* elements can be inferred based on rice:barley synteny, as it has been shown that the genomes of cereal grasses share colinearity in gene order in most cases, as demonstrated by the hybridization of probes to specific DNA sequences from various members of the grass family (Gale and Devos 1998). Preliminary studies have assessed the syntenic relationships of QTLs for disease resistance in rice and barley (Chen et al. 2003) and compared the colinearity of barley and rice genes located on selected BACs (Dubcovsky et al. 2001). In order for this to be effective as a tool for mapping the genomic DNA sequences flanking the *Ds* insertions, many more markers from barley need to be mapped in the rice genome, and a more detailed understanding of the syntenic relationship of specific genes and gene regions in rice and barley will be needed.

The approximate locations of the *Ds* insertions could also be determined using wheat deletion lines (Qi et al. 2003) or the wheat: barley addition lines (Islam et al. 1981). Wheat deletion line mapping could potentially provide resolution in the 10-Mb range (http://wheat.pw.usda.gov/NSF/progress_mapping.html), but the insertion events would not be positioned on a barley linkage map. With wheat:barley addition lines the *Ds* insertions could only be located on a chromosome arm, or a sector of a chromosome arm. These approaches, in general, lack the resolution necessary to assign *Ds* elements to specific barley linkage map bins.

Two other approaches that were considered for mapping of the *Ds* insertions were bulked segregant analysis (Michelmore et al. 1991) in the progeny of each TNP line and OWB-D, and RFLP analysis. We opted for the sequence-based strategy in order to develop a better understanding of the architecture of the barley genome and the sequence context of each *DsT* locus. We utilized RFLP mapping to validate the sequence-based mapping approach, but decided not to use it as our primary strategy due to the substantially large amounts of genomic DNA and radioactivity required for efficient use of this technique.

Characterization of insertion site sequences

We were interested in characterizing the insertion sites of the *Ds* elements, since it has been estimated that genes only account for ~11% of the barley genome, while long terminal repeat retrotransposons and other repetitive elements represent more than 51% of the genome (Rostoks et al. 2002). The Morex BAC library consists of 313,344 clones on 17 filters, providing approximately 6.3 haploid genome equivalents (Yu et al. 2000). Preliminary screening with TNP flanking sequences probes utilized a subset of the 17 filters, representing approximately 1.5-fold (four filters) or 3.0-fold (8 filters) coverage of the genome. Yu et al. (2000) screened the entire BAC library with 40 single-copy probes and found an average of 6.4 positive BAC clones per probe, with a range of 1–13 clones per probe. Most *Ds* flanking regions probes were of low copy number, as evidenced by their hybridization to between one and 10 BAC clones (Table 2; see Electronic Supplementary Material for addresses of positive BAC clones). Hybridization of mapped and cloned *Ds* flanking sequences to the Morex genomic BAC library supported our estimates of copy number based on RFLP analysis and BLAST search results. These results support the hypothesis that, in barley, as observed in other heterologous plant systems such as *Arabidopsis* (Parinov et al. 1999; Tissier et al. 1999; Ito et al. 2002) and rice (Enoki et al. 1999; Greco et al. 2001; Kolesnik et al. 2004), the *Ds* element transposes preferentially into predicted coding regions.

There is evidence that some of the *Ds* insertions reside in members of multigene families. The TNP-1 probe hybridized strongly to seven clones on four of 17 filters, but slightly weaker hybridization was apparent to another 10 clones, suggesting that similar, but not identical, sequences occur at multiple sites in the barley genome. The DNA sequence flanking this *DsT* locus is similar to an uncharacterized EST from maize, based on BLAST search results (Table 2). There were two other examples of *Ds* insertions into members of multigene families: in one case a putative cytochrome P450 (*OSU-STS1*) was tagged, and in another an uncharacterized sequence *(OSU-STS 2* and *3*). In both cases, we were able to map the two members of the family to different chromosomal locations, but we were unable to resolve which of the gene family members contained the *Ds* insertion. Thus, these sites were added onto the map as sequence tagged sites only, rather than ambiguous insertion sites.

Comparison of the flanking sequences using BLASTN and TBLASTX searches in various publicly available EST and nucleotide databases also supports the hypothesis that *Ds* tends to insert into predicted coding regions in the genome (Table 2). The best match found in the BLAST search is to a member of a well-characterized multigene family, the wall-associated protein kinases (TNP-11). In the line TNP-22, the *Ds* element has inserted into a putative aristolochene synthase gene, encoding a terpenoid biosynthesis enzyme. In total, eleven of the nineteen insertions, or about 65%, are located in predicted or characterized genes, based upon their similarity to ESTs from members of the Triticeae or rice. This rate is similar to that seen in studies of the *Ac*/*Ds* in rice, where Kolesnick et al. (2004) estimated that about 72% of the *Ds* insertions (on chromosome 1) were in genic regions, while Greco et al. (2001) found that approximately 66% of the *Ds* insertion sites showed homology to predicted genes in public databases. Two *Ds* insertions (which were not mapped) have been localized in suspected retrotransposon sequences, identified by comparison to the TIGR Plant Repeat database for *Hordeum*. One insertion is located in a *copia* -like BARE-1 retrotransposon ($1.1e^{-87}$) and the other is in a *Hordeum vulgare* Bagy-2 retrotransposon ($8.4e^{-19}$). In some cases (*DsT-27, 28, 30, 31, 33, 34*) the BLAST searches resulted in Expectation (E) values below what we considered sufficient ($1e^{-4}$) to assign a putative identity to the flanking sequences. These sequences may represent previously unidentified genes, but the BAC hybridization evidence indicates that these *DsT* loci do not reside in repetitive DNA sequences.

Conclusion

Our principal objective was to map *Ds* insertions in the barley genome (using primarily the OWB DH population) and assign them bin map positions to facilitate integration with other genomics resources in barley. We were interested in assessing the effectiveness of the sequence-based mapping strategy and performing a preliminary characterization of the sites of insertion of the *Ds* elements. Our sequence-based approach for mapping *Ds* insertions in barley proved to be effective; all *DsT* loci were assigned to linkage map positions, and their positions were confirmed by mapping a subset of four of the same insertions by RFLP analysis. These data indicate that *Ds* insertions are located throughout the genome, in 17 bins, on six of the seven chromosomes. BLAST searches and screens of the barley BAC library using cloned flanking DNA sequences indicated that the majority of the *Ds* insertions reside in predicted coding regions of the DNA. The development of this transposon-tagging resource in barley promises to provide a useful tool for functional genomics studies in Triticeae.

Acknowledgements This work was supported by NSF Award No. 0110512. PGL is supported by USDA Cooperative Extension through the University of California

References

Altschul S, Gish W, Miller W, Myers E, Lipman D (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Baker B, Schell J, Lorz H, Fedoroff N (1986) Transposition of the maize controlling element ' *Activator* ' in tobacco. Proc Natl Acad Sci USA 83:4844–4848

Batley J, Barker G, O'Sullivan H, Edwards KJ, Edwards D (2003) Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. Plant Physiol 132:84–91

Bennett M, Smith J (1976) Nuclear DNA amounts in angiosperms. Phil Trans Roy Soc Lond Series B 274:227–274

Chen H, Wang S, Xing Y, Xu C, Hayes PM, Zhang Q (2003) Comparative analyses of genomic locations and race specificities of loci for quantitative resistance to *Pyricularia grisea* in rice and barley. Proc Natl Acad Sci USA 100:2544–2549

Costa JM, Corey A, Hayes P, Jobet C, Kleinhofs A, Kopisch-Obusch A, Kramer SF, Kudrna D, Li M, Riera-Lizarazu O, Sato K, Szucs P, Toojinda T, Vales MI, Wolfe RI (2001) Molecular mapping of the Oregon Wolfe barleys: a phenotypically polymorphic doubled-haploid population. Theor Appl Genet 103:415–424

Dellaporta S (1994) Plant DNA miniprep and microprep. In: Freeling M, Walbot V (eds) The Maize Handbook. Springer-Verlag, New York, pp 522–525

Dubcovsky J, Ramakrishna W, SanMiguel PJ, Busso CS, Yan L, Shiloff BA, Bennetzen JL (2001) Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. Plant Physiol 125:1342–1353

Enoki H, Izawa T, Kawahara M, Komatsu M, Koh S, Kyozuka J, Shimamoto K (1999) *Ac* as a tool for the functional genomics of rice. Plant J 19:605–613

Fedoroff N, Wessler S, Shure M (1983) Isolation of the transposable maize controlling elements *Ac* and *Ds*. Cell 35:235–242

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Feinberg AP, Vogelstein B (1984) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity, Addendum. Anal Biochem 137:266–267

Gale M, Devos K (1998) Comparative genetics in the grasses. Proc Natl Acad Sci USA 95:1971–1974

Greco R, Ouwerkerk PBF, Sallaud C, Kohli A, Colombo L, Puigdomenech P, Guiderdoni E, Christou P, Hoge JHC, Pereira A (2001) Transposon insertional mutagenesis in rice. Plant Physiol 125:1175–1177

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hayes PM, Castro A, Marquez-Cedillo L, Corey A, Henson C, Jones BL, Kling J, Mather D, Matus I, Rossi C, Sato K (2003) Genetic diversity for quantitatively inherited agronomic and malting quality traits. In: von Bothmer R, Knüpffer H, van Hintum T, Sato K (eds) Diversity in barley (*Hordeum vulgare*). Elsevier, Amsterdam pp 201–226

Islam AKMR, Shepherd KW, Sparrow DHB (1981) Isolation and characterization of euplasmic wheat-barley chromosome addition lines. Heredity 46:161–174

Ito T, Motohashi R, Kuromori T, Mizukado S, Sakurai T, Kanahara H, Seki M, Shinozaki K (2002) A new resource of locally transposed *Dissociation* elements for screening gene-knockout lines *in silico* on the *Arabidopsis* genome. Plant Physiol 129:1695–1699

Jähne A, Becker D, Brettschneider R, Lörz H (1994) Regeneration of transgenic, microspore-derived, fertile barley. Theor Appl Genet 89:525–533

Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH (2000) Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. Proc Natl Acad Sci USA 97:6603–6607

Kleinhofs A, Graner A (2001) An integrated map of the barley genome. In: Phillips RL, Vasil IK (eds) DNA-based markers in plants. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 187–199

Kleinhofs A, Kilian A, Saghai Maroof MA, Biyashev RM, Hayes P, Chen FQ, Lapitan N, Fenwick A, Blake TK, Kanazin V (1993) A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. Theor Appl Genet 86:705–712

Kolesnik T, Szeverenyi I, Bachmann D, Kumar CS, Jiang S, Ramamoorthy R, Cai M, Ma ZG, Sundaresan V, Ramachandran S (2004) Establishing an efficient *Ac*/*Ds* tagging system in rice: large-scale analysis of *Ds* flanking sequences. Plant J 37:301–314

Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4:403–410

Koprek TK, McElroy D, Louwerse J, Williams-Carrier R, Lemaux PG (2000) An efficient method for dispersing *Ds* elements in the barley genome as a tool for determining gene function. Plant J 24:253–263

Koprek T, Rangel S, McElroy D, Louwerse JD, Williams-Carrier RE, Lemaux PG (2001) Transposon-mediated single-copy gene delivery leads to increased transgene expression stability in barley. Plant Physiol 125:1354–1362

Kota R, Rudd S, Facius A, Kolesov G, Thiel T, Zhang H, Stein N, Mayer K, Graner A (2003) Snipping polymorphisms from large EST collections in barley (*Hordeum vulgare* L.). Mol Genet Genomics 270:24–32

Li J, Lease KA, Tax FE, Walker JC (2001) BRS1, a serine carboxypeptidase, regulates, BRI1 signaling in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 98:5916–5921

Liu Y-G, Whittier RF (1994) Preparation of megabase plant DNA from nuclei in agarose plugs and microbeads. Nucleic Acids Res 22:2168–2169

Liu Y-G, Mitsukawa N, Oosumi T, Whittier RF (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J 8:457–463

McElroy D, Louwerse J, McElroy S, Lemaux PG (1997) Development of a simple transient assay for *Ac*/*Dc* activity in cells of intact barley tissue. Plant J 11:157–165

Mgonja MA, Dahleen LS, Franckowiack JD (1995) Subsets from mapping populations for localization of new genes in barley. Barley Genet News 24:14–23

Michelmore R, Paran I, Kesseli R (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Moore G, Devos K, Wang Z, Gale M (1995) Cereal genome evolution. Grasses, line up and form a circle. Curr Biol 5:737–739

Ochman H, Gerber AS, Hartl DL (1988) Genetic applications of an inverse polymerase chain reaction. Genetics 120:621–623

Ochman H, Ayala FJ, Hartl DL (1993) Use of polymerase chain reaction to amplify segments outside boundaries of known sequences. Methods Enzymol 218:309–321

Pan A, Hayes PM, Chen THH, Blake TK, Wright S, Karsai I, Bedö Z (1994) Genetic analysis of the components of winter hardiness in barley (*Hordeum vulgare* L.). Theor Appl Genet 89:900–910

Parinov S, Sevugan M, Ye D, Yang W-C, Kumaran M, Sundaresan V (1999) Analysis of flanking sequences from *Dissociation* insertion lines: a database for reverse genetics in *Arabidopsis*. Plant Cell 11:2263–2270

Qi L, Echalier B, Friebe B, Gill BS (2003) Molecular characterization of a set of wheat deletion stocks for use in chromosome bin mapping of ESTs. Funct Integr Genomics 3:39–55

Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5:94–100

Raina S, Mahalingam R, Chen F, Fedoroff N (2002) A collection of sequenced and mapped *Ds* transposon insertion sites in *Arabidopsis thaliana*. Plant Mol Biol 50:91–108

Ramsay L, et al (2000) A simple sequence repeat-based linkage map of barley. Genetics 156:1997–2005

Rostoks N, Park Y-J, Ramakrishna W, Ma J, Druka A, Shiloff BA, SanMiguel PJ, Jiang Z, Brueggeman R, Sandhu D, Gill K, Bennetzen JL, Kleinhofs A (2002) Genomic sequencing reveals gene content, genomic organization, and recombination relationships in barley. Funct Integr Genomics 2:51–59

Sambrook J, Russell DW. 2001. Molecular cloning: a laboratory manual (3rd edn). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Scholz S, Lörz H, Lütticke S (2001) Transposition of the maize transposable element *Ac* in barley (*Hordeum vulgare* L.). Mol Gen Genet 264:653–661

Thiel T, Michalek W, Varshney RK, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor Appl Genet 106:411–422

Tissier AF, Marillonnet S, Klimyuk V, Patel K, Torres MA, Murphy G, Jones JDG (1999) Multiple independent defective *Suppressor-mutator* transposon insertions in *Arabidopsis*: a tool for functional genomics. Plant Cell 11:1841–1852

Van Ooijen JW, Voorrips RE (2001) JoinMap 3.0, Software for the calculation of genetic linkage maps. Release 3.0. Plant Research International, Wageningen, The Netherlands

Wan Y, Lemaux PC (1994) Generation of large numbers of independently transformed fertile barley plants. Plant Physiol 104:37–48

Wanamaker S, Close T (2003) HarvEST: Triticeae. Release 0.99 (available at http://138.23.191.152:/blast/blast.html)

Waugh R, McLean K, Flavell A, Pearce S, Kumar A, Thomas BBT, Powell W (1997) Genetic distribution of BARE-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Mol Gen Genet 253:687–694

Yu Y, Tomkins JP, Waugh R, Frisch DA, Kudrna D, Kleinhofs A, Brueggeman RS, Muehlbauer GJ, Wise RP, Wing RA (2000) A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. Theor Appl Genet 101:1093–1099